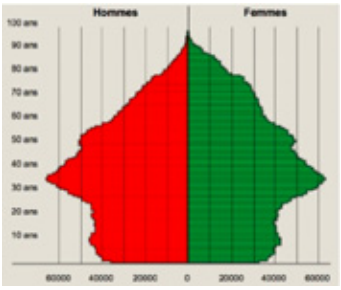
Thème 20: Quelques éléments de statistique descriptive

§20.1 Introduction, un peu d'histoire:

Introduction

Le mot statistique – de l'italien « statista », homme d'État – désignait à l'origine la collecte et l'évaluation des données concernant un État. Cette science de l'État était une représentation purement descriptive de faits géographiques et sociaux comme le climat, la population, les coutumes, les organisations économiques, etc…, à l'usage des hommes d'État ; à l'époque en France le roi et son conseil.



Pyramide des âges
Suisse 2000

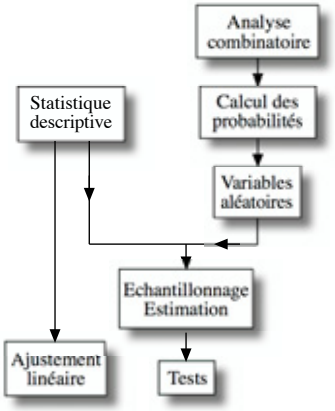
Dès la plus haute Antiquité, les dirigeants ont fait procéder à des enquêtes sur la population: l'Empereur Yao (vers 2200 av. J.-C.) pour connaître les productions agricoles, les pharaons égyptiens (dès 1700 av. J.-C.), l'Empereur Auguste à Rome pour le nombre de soldats, les revenus des citoyens.

Une page des données de l'astronome Tycho Brahé

Nous trouvons également de multiples exemples d'utilisation de statistiques dans les sciences :

- *Johannes Kepler (1571-1630) formula ses lois sur les mouvements des planètes en utilisant l'ensemble des données récoltées par l'astronome danois Tycho Brahé (1546-1601).*
- *Les premières études statistiques de Florence Nightingale, infirmière anglaise durant la guerre de Crimée de 1854 à 1856. permirent d'identifier les causes de mortalité des soldats et conduisirent à l'amélioration des conditions d'hygiène des hôpitaux militaires anglais.*

Aujourd'hui, cette partie des mathématiques a pris une grande place grâce aux nouvelles techniques et à la puissance des ordinateurs. Géographie, médecine, sciences humaines, sciences économiques, biologie, politique, aucun domaine n'est épargné.



On peut décomposer la méthode statistique en cinq étapes:

1. Identification précise de la population et du (des) caractère(s) à étudier
2. Récolte des données (recensement ou échantillonnage)
3. Regroupement, classification et présentation des données (statistiques descriptives)
4. Comparaison avec des modèles théoriques (calcul des probabilités et modèles probabilistes)
5. Interprétation, conclusion, prévision (inférence statistique)

§20.2 Vocabulaire:

En statistique, le mot **population** représente un ensemble d'objets de même nature que l'on va étudier, analyser. Les éléments de la population, appelés **individus**, peuvent être des personnes, mais aussi des choses, des animaux, des objets, des faits, des notes de TE, etc... Le nombre d'individus est appelé l'**effectif**.

*Souvent, il n'est pas possible de prendre en compte la totalité de la population. Dans ces cas, l'étude se limite à un **échantillon**, pris au hasard, à partir duquel on peut tenter de déduire une tendance pour toute la population.*

Une population doit toujours être clairement définie afin que l'on puisse toujours déterminer si un élément quelconque fait ou non partie de la population étudiée. On pourra ainsi étudier une caractéristique que possède chacun des individus on appelle cela une **variable statistique (v.s)**.

Les différentes valeurs que peut prendre une variable statistique sont les **modalités** de cette variable.

Notation : On note une v.s par une lettre majuscule X, Y, … et ses modalités par la même lettre minuscule affectée d'indices : x_1, x_2, … pour la variable X ou y_1, y_2, … pour la variable Y.

Modèle 1 : On fait une étude statistique auprès des élèves du gymnase de Morges. On aimerait connaître le sexe, l'âge au 1er janvier, la taille, la voie (ECGC ou EM) de chaque élève.

Population :

v.s	modalité des v.s
X :	x_1 = x_2 =
Y :	y_1 = y_2 = y_3 = y_4 = y_5 =
Z :	$z_i \in$ [;]
U :	u_1 = u_2 =

Une v.s. est **quantitative** si les valeurs qu'elle peut prendre sont **numériques**. Une telle v.s est dite **quantitative discrète** si les valeurs qu'elle peut prendre sont isolées les unes des autres. Par contre, si celles-ci constituent des intervalles de nombres, la v.s est appelée **quantitative continue**. Si les valeurs d'une v.s sont descriptives ou nominatives, la v.s. est dite **qualitative**.

X est une v.s ……, Y est une v.s ……

Z est une v.s ……, U est une v.s ……

Exercice 20.1: On a demandé aux employés d'une entreprise pour quel parti politique ils avaient voté lors des dernières élections. Voici les données brutes obtenues:

PS	PRD	PS	PDC	PS	UDC
PS	UDC	PRD	PS	verts	PDC
UDC	PRD	verts	UDC	UDC	UDC
PRD	PS	PRD	PDC	PRD	PDC
UDC	PDC	PS	UDC	UDC	UDC

a) Identifier la population ainsi que la variable statistique (v.s.).
b) Donner l'ensemble des modalités.
c) De quel type est cette variable statistique ?

Exercice 20.2: Un professeur de l'Uni a noté le nombre de points obtenus par 80 étudiants lors d'un test de statistiques.

2	3	5	5	4	6	6	5	4	3
7	7	7	6	2	7	7	9	8	10
5	6	6	8	6	6	3	7	3	5
9	7	6	4	7	5	9	9	6	9
6	3	9	8	8	7	5	6	10	6
9	7	7	7	4	7	10	8	7	10
3	5	8	5	8	7	4	8	10	7
4	6	6	8	7	7	7	8	8	9

a) Identifier la population ainsi que la variable statistique (v.s.).
b) Donner l'ensemble des modalités.
c) De quel type est cette variable statistique ?

Modèle 2 : En reprenant les données de l'exercice 20.1, on va sacrifier le caractère individuel de l'information afin d'obtenir un portrait d'ensemble. On calcule pour chaque modalité le nombre d'individus ayant cette modalité : l'**effectif n_i** de la modalité. Celui-ci ne permet pas de comparer deux populations inégales. Il sera alors naturel de calculer la proportion de la population qui a une telle modalité. On définit alors la **fréquence relative f_i** par le rapport entre l'effectif de chaque modalité et le nombre N d'individus de la population: $f_i = \frac{n_i}{N}$

Modalité x_i	**Effectif n_i**	**Fréquence relative f_i**	**Angle**
PS			
PRD			
PDC			
UDC			
Verts			
Total:			

Le **tableau de distribution des effectifs et des fréquences** permet une bonne synthèse des informations, mais n'est pas très explicite. On l'accompagnera d'un graphique permettant de représenter ces données. On utilise fréquemment :

a) un diagramme en colonnes (histogramme)

b) un diagramme en secteurs (en "camembert")

Remarques :

- La somme des effectifs est toujours égale au nombre d'individus de la population:

$$n_1 + n_2 + \ldots + n_k = N$$

- La somme des fréquences est toujours égale à 1:

$$f_1 + f_2 + \ldots + f_k = 1$$

car:

Lors de l'utilisation de la calculatrice, il peut arriver que la somme des fréquences ne soit pas exactement égale à 1 à cause des arrondis de calculs.

Exercice 20.3: Reprendre les données de l'exercice 20.2 afin d'en proposer :

a) le tableau de distribution des effectifs et des fréquences ;

b) un histogramme puis un diagramme en secteurs.

Exercice 20.4: On a demandé aux enfants de trois classes de 3ème année primaire quel était leur sport d'hiver préféré. On a obtenu les données brutes suivantes:

Hockey	Glissade	Hockey	Hockey	Hockey
Hockey	Ski	Hockey	Ski	Raquette
Patinage	Ski	Ski	Hockey	Ski
Ski	Hockey	Ski	Raquette	Ski
Patinage	Ski	Hockey	Raquette	Raquette
Ski	Glissade	Hockey	Glissade	Glissade
Hockey	Glissade	Hockey	Hockey	Hockey
Ski de fond	Hockey	Patinage	Patinage	Hockey
Ski	Hockey	Ski	Raquette	Patinage
Hockey	Glissade	Ski	Ski	Ski de fond
Hockey	Patinage	Ski	Patinage	Hockey
Hockey	Patinage	Ski	Patinage	Raquette

a) Identifier la population.
b) Caractériser la variable statistique.
c) Donner l'ensemble des modalités.
d) Le tableau des distributions des effectifs et des fréquences.
e) Faire un diagramme en secteurs.

Exercice 20.5: On étudie l'état civil des 30 employés (numérotés de 1 à 30) d'une petite entreprise.

1	Marié	11	Marié	21	Célibataire
2	Mariée	12	Célibataire	22	Marié
3	Célibataire	13	Marié	23	Veuf
4	Divorcé	14	Veuve	24	Célibataire
5	Marié	15	Marié	25	Divorcée
6	Célibataire	16	Divorcé	26	Divorcé
7	Célibataire	17	Célibataire	27	Marié
8	Mariée	18	Mariée	28	Marié
9	Mariée	19	Marié	29	Marié
10	Divorcée	20	Marié	30	Marié

a) Identifier la population
b) Caractériser la variable statistique
c) Donner l'ensemble des modalités
d) Le tableau des distributions des effectifs et des fréquences
e) Proposer l'histogramme **des effectifs** de cette v.s.
f) Proposer l'histogramme **des fréquences** de cette v.s.
g) Comparer ces 2 représentations graphiques

§20.3 Regroupement des données à l'intérieur de classes de valeurs

Souvent, lors d'une étude statistique portant sur une variable statistique quantitative discrète ou continue, les données recueillies diffèrent à peu près toutes les unes des autres et sont étalées sur un large intervalle de valeurs. L'objectif de la statistique descriptive étant de résumer de la façon la plus adéquate possible cet ensemble de données, les mesures seront regroupées dans des intervalles de valeurs que l'on appelle des classes du type $[b_{i-1} ; b_i[$.

b_{i-1} est **la borne inférieure** de la classe i;

b_i est **la borne supérieure** de la classe i;

$x_i = \frac{b_{i-1} + b_i}{2}$ est **le milieu** de la classe i;

$L_i = b_i - b_{i-1}$ est **la largeur** de la classe i.

Modèle 3 : Des chimistes viennent de composer une nouvelle fibre synthétique qui devrait se caractériser par sa résistance. Afin de vérifier sa capacité de tension, on prélève de la production, au hasard, un échantillon de 60 fibres qu'on soumet à des essais de résistance. Les résultats (en kg) sont les suivants :

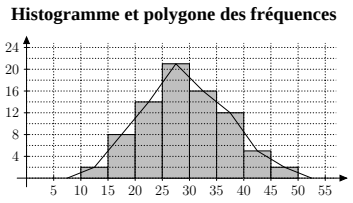
35	65	71	75	77	80	81	82	84	86	87	89	91	97	100
48	69	72	75	78	80	81	83	85	86	88	89	94	97	103
53	69	73	76	79	80	81	83	85	87	88	89	95	99	104
63	71	74	77	79	81	82	84	86	87	89	91	97	99	114

a) On regroupe les données en 6 classes d'amplitude 15 avec 30 comme valeur minimale. Compléter le tableau:

Classe	Centre	Effectif	Fréquence
[30 ; 45[			
[45 ; 60[			
[60 ; 75[			
[75 ; 90[			
[90 ; 105[			
[105 ; 120[			
Totaux			

Représentations graphiques:

- L'**histogramme** est un diagramme en colonnes où les rectangles sont juxtaposés indiquant ainsi le caractère continu de la variable statistique.



- Le **polygone des fréquences** est la ligne polygonale obtenue en joignant les points milieux consécutifs des sommets des rectangles de l'histogramme. On commence et on termine le polygone des fréquences en ajoutant une classe de fréquence nulle avant la première classe et une autre après la dernière classe.

Modèle 3 Suite : À propos de cette nouvelle fibre synthétique:

b) Effectuer l'histogramme.
c) Construire le polygone des fréquences.

Exercice 20.6: Une entreprise a enregistré le salaire de tous ses vendeurs pour l'année dernière. Voici les données rangées:

Classes (salaires)	x_i	n_i	f_i
[10000 ; 15000[	12500	2	……%
[15000 ; 20000[	……	8	10,00%
[20000 ; 25000[	22500	14	……%
[25000 ; 30000[	27500	21	26,25%
[30000 ; 35000[	……	……	20,00%
[35000 ; 40000[	37500	12	15,00%
[40000 ; 45000[	42500	5	6,25%
[45000 ; 50000[	47500	……	2,50%
	Totaux	80	100

a) Compléter le tableau des distributions des effectifs et des fréquences.
b) Faire un histogramme.
c) Construire le polygone des fréquences.

Exercice 20.7: En recevant les élèves qui désirent faire partie d'une équipe de foot du gymnase, l'entraîneur a pris note du poids de ces 60 joueurs:

72,6	81,9	84,7	88,1	89,4	91,6	93,7	95,8	99,1	103,2
75,8	82,6	85,4	88,1	90,2	92,4	93,9	96,6	99,4	103,9
77,5	82,9	86,2	88,3	90,9	92,5	94,4	97,1	99,8	104,0
78,3	83,0	86,9	88,7	91,1	92,8	94,7	97,2	100,4	105,2
79,6	83,5	87,3	89,0	91,2	93,0	94,8	97,5	101,7	106,1
81,5	84,1	87,8	89,1	91,3	93,3	95,2	98,3	102,1	118,7

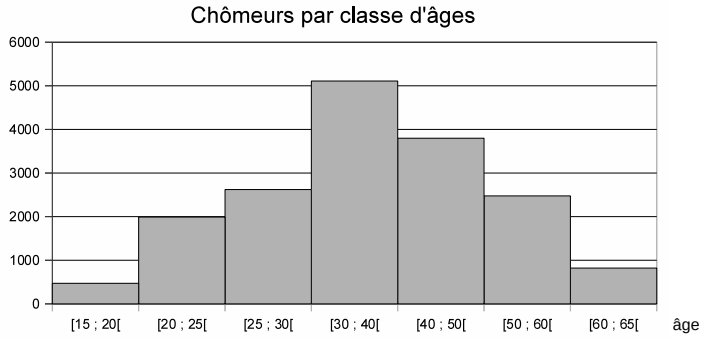
a) Identifier la population.
b) Identifier la variable statistique.
c) Cette variable statistique est-elle discrète ou continue ?
d) En utilisant des classes de largeur 5, construire le tableau des distributions des effectifs et des fréquences (valeur minimale: 70). Vous admettrez une classe plus large à l'extrémité (classe [105 ; 120[ ),
f) Construire le polygone des fréquences.

Exercice 20.8: Le tableau récapitulatif suivant donne la statistique trimestrielle par classe d'âges des chômeurs inscrits dans un office du travail dans le canton de Vaud en juin 2009:

Chômeurs par classe d'âges

Classe d'âges	Effectif	Fréquence [%]
[15 ; 20[	472	2,73
[20 ; 25[	1'990	11,51
[25 ; 30[	2'621	15,16
[30 ; 40[	5'110	29,56
[40 ; 50[	3'798	21,97
[50 ; 60[	2'476	14,32
[60 ; 65[	821	4,75
Totaux	17'288	100

a) En quoi l'histogramme suivant est-il trompeur ?



b) Proposer un nouvel histogramme corrigeant cet effet visuel trompeur.

§20.4 Fréquences cumulées des v.s. quantitatives

Dans une étude statistique, si on souhaite connaître la proportion de chaque valeur que peut prendre la variable statistique étudiée, on regarde sa **fréquence f_i**.

Si par contre on souhaite connaître la proportion des individus qui présentent des valeurs inférieures ou égales à une valeur fixée, on regarde la **fréquence cumulée croissante F_i**.

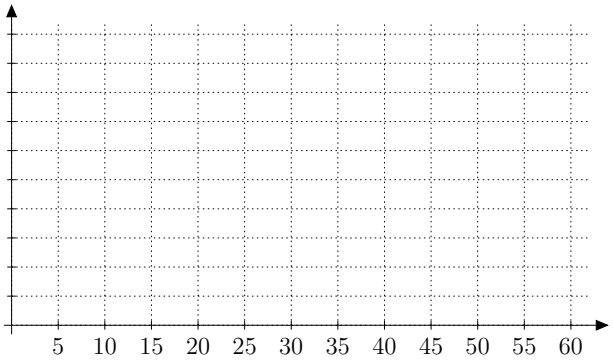
Observons ceci sur le modèle suivant :

Modèle 4 : Lors d'un concours de pêche dans le lac de Bret, on a mesuré (en cm) toutes les prises et regroupées par classe dans le tableau qui suit :

a) Compléter le tableau que l'on a obtenu :

Classe $[b_{i-1} ; b_i [$	**Centre** x_i	**Effectifs** n_i	**Fréquence** f_i	**Fréquence cumulée croiss** F_i
[10 ; 15[	……	4	……	6,25 %
[15 ; 20[	……	8	12,5 %	……
[20 ; 25[	22,5	……	32,81 %	51,56 %
[25 ; 30[	27,5	18	28,13 %	79,69 %
[30 ; 35[	32,5	7	10,94 %	……
[35 ; 40[	37,5	5	……	98,44
[40 ; 60[	50	1	1,56 %	……
	Totaux :	64	100 %	

b) Représenter le diagramme des fréquences cumulées.



c) Déterminer la proportion des prises dont la longueur est plus petite que 30 cm.

d) Déterminer la proportion des prises dont la longueur est plus grande ou égale à 20 cm.

La médiane : L'abscisse du point correspondant à une fréquence cumulée croissante de 50% s'appelle **la médiane** de la v.s.

Dans le modèle précédent, la médiane vaut:

Exercice 20.9: Compléter la solution de l'exercice 20.8 par:

c) Représenter les courbes des fréquences cumulées croissantes et en déduire la valeur de la médiane.
d) Déterminer la proportion de chômeurs de moins de 40 ans.
e) Déterminer la proportion de chômeurs de plus de 20 ans.
f) Déterminer la proportion des chômeurs entre 20 et 40 ans.

Exercice 20.10: Lors d'un cours de statistique, en 2008, 20 étudiants ont été invités à indiquer leur taille et leur poids.

N° d'ordre	taille en cm.	poids en kg
1	174	64
2	175	59
3	180	64
4	168	62
5	175	50
6	170	60
7	170	68
8	160	63
9	187	93
10	178	70

N° d'ordre	taille en cm.	poids en kg.
11	170	64
12	182	72
13	168	60
14	171	55
15	181	80
16	178	82
17	180	72
18	180	78
19	178	71
20	182	72

a) Regrouper les données de tailles et de poids en 6 classes de largeur égale.
b) Représenter l'histogramme et le polygone des fréquences des tailles des 20 étudiants.
c) le diagramme des fréquences cumulées croissantes des poids des 20 étudiants.

§20.5 Les mesures de tendance centrale d'une variable discrète (*k* modalités)

Introduction Nous avons vu dans les paragraphes précédents comment résumer un grand nombre de données sous la forme de tableaux ou de diagrammes. Il est pourtant souvent possible de caractériser une distribution de manière beaucoup plus succincte par une mesure de l' "emplacement" du centre et une mesure de la dispersion des observations autour de ce centre.

Dans ce chapitre, nous examinerons la première des deux caractéristiques d'une v.s quantitative soit les **mesures de tendance centrale**. Nous distinguerons deux types de mesure relative au centre de la distribution qui sont utilisés les plus fréquemment: **la moyenne, la médiane**.

La moyenne arithmétique $\overline{x}$:

$$\overline{x} = \frac{n_1 \cdot x_1 + n_2 \cdot x_2 + \ldots + n_k \cdot x_k}{n_1 + n_2 + \ldots + n_k}$$

où sous la forme:

$$\overline{x} = f_1 x_1 + f_2 x_2 + \ldots + f_k x_k$$

La médiane *M*: La médiane *M* d'une variable discrète est la première modalité dont la fréquence cumulée croissante atteint ou dépasse 50%.

Modèle 5: Considérons le nombre de personnes par ménage dans le canton de Neuchâtel en 1980.

x_i	n_i			
1	20'734			
2	20'798			
3	10'067			
4	10'381			
5	3'053			
6	832			
Totaux:				

Dans ce tableau, nous avons $x_1 = 1$, $x_2 = 2$, …, les x_i représentent le nombre de personnes par ménage. $n_1 = 20'734$, $n_2 = 20'798$,…, les n_i indiquant le nombre de ménages comportant x_i personnes.

Calculons les mesures de tendance centrale de cette distribution.

Exercice 20.11: Calculer la moyenne, la médiane de la v.s suivante:

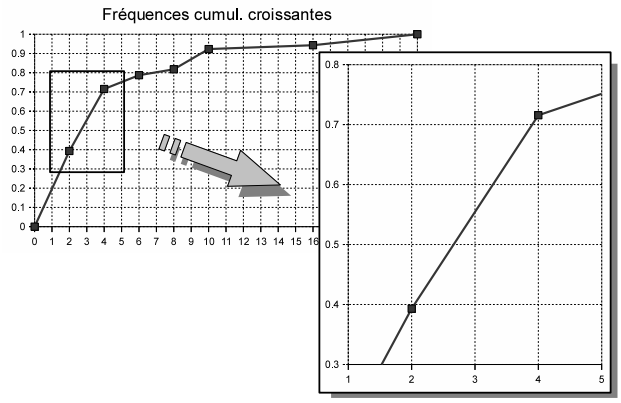
Modalités	Effectifs			
10	2			
11	3			
12	7			
13	9			
14	14			
15	8			
16	3			
17	1			

§20.6 Les mesures de tendance centrale d'une variable continue

La moyenne arithmétique $\bar{x}$: La moyenne arithmétique $\bar{x}$ d'une variable statistique continue est calculée comme si toutes les données étaient situées au centre de la classe. On retrouve donc:

$$\bar{x} = f_1 x_1 + f_2 x_2 + \ldots + f_k x_k$$

La médiane *M*: Comme déjà mentionné préalablement, la médiane se calcule en utilisant le polygone des fréquences cumulées et le théorème de Thalès:



$$M = b_{i-1} + \frac{0{,}50 - F_{i-1}}{f_i} \cdot L_i$$

avec b_{i-1} la borne inférieure de la 1ère classe dont la fréquence cumulée croissante atteint ou dépasse 50%;
f_i la fréquence de cette 1ère classe ;
F_{i-1} la fréquence cumulée croissante de la classe précédente;
L_i la largeur de la classe.

Modèle 6 : On considère la v.s continue donnée dans le tableau suivant:

$[b_{i-1}\ ;\ b_i[$	n_i	f_i	x_i	$f_i x_i$	F_i
[30 ; 40[	4				
[40 ; 50[	7				
[50 ; 60[	11				
[60 ; 70[	12				
[70 ; 80[	8				
[80 ; 90[	5				
Totaux	47				

Déterminer la moyenne et la médiane de cette v.s.

Exercice 20.12: Calculer la médiane de la v.s continue suivante:

$[b_{i-1}\ ;\ b_i[$	n_i
[0 ; 2[	3
[2 ; 4[	8
[4 ; 6[	15
[6 ; 8[	14
[8 ; 10[	6
[10 ; 12[	2
Totaux	

Exercice 20.13: Le club PAD organise un grand tournoi de quilles. Voici le tableau de distribution des scores:

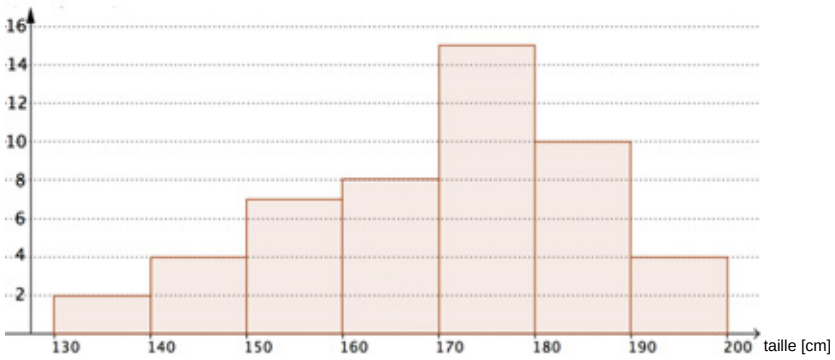
$[b_{i-1}\ ;\ b_i[$	n_i				
[120 ; 140[	1				
[140 ; 160[	9				
[160 ; 180[	22				
[180 ; 200[	51				
[200 ; 220[	12				
[220 ; 240[	5				
Totaux	100				

Déterminer la moyenne et la médiane des scores obtenus.

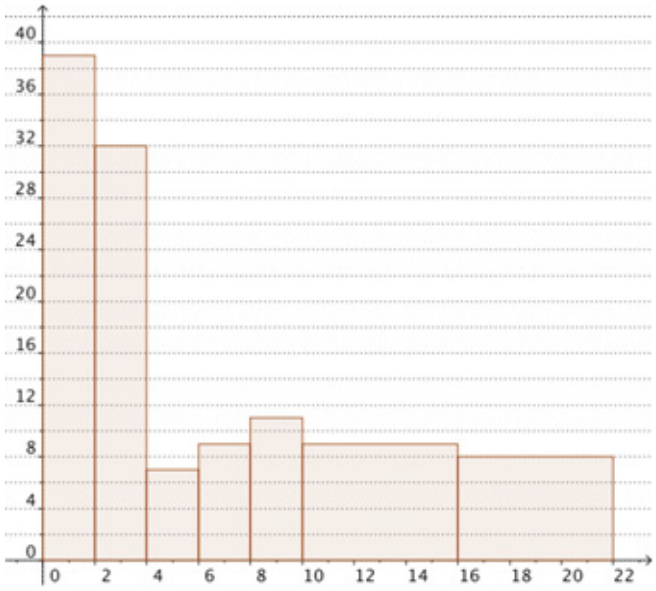
Exercice 20.14: On a mesuré la taille des 50 professeurs du gymnase de Morges:

Taille en cm	Nombre de professeurs
[130 ; 140[	2
[140 ; 150[	4
[150 ; 160[	7
[160 ; 170[	8
[170 ; 180[	15
[180 ; 190[	10
[190 ; 200[	4

Calculer la moyenne et la médiane puis les représenter sur l'histogramme suivant:

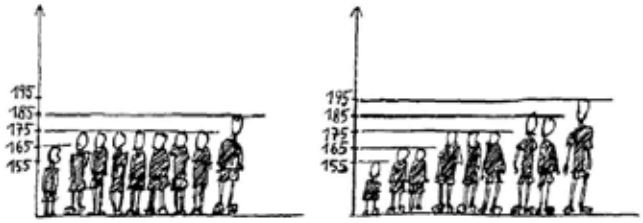


Exercice 20.15: La médiane d'une variable statistique continue peut également être interprétée comme la valeur qui divise l'histogramme en deux parties de la même aire.
Déterminer graphiquement sur l'histogramme suivant la médiane de cette v.s.



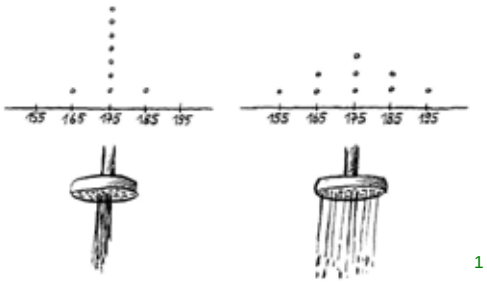
§20.7 Les mesures de dispersion d'une variable statistique

Introduction Le paragraphe précédent a été consacré à l'étude de deux mesures de tendance centrale. Elles indiquent autour de quelle valeur se situent les données, mais ne donnent pas une description suffisante de la variable statistique. Par exemple, si on désire comparer les 2 groupes d'élèves proposés dans les diagrammes ci-dessous:



$\overline{x}$ = …… $\overline{x}$ = ……

Mais pourtant, les 2 distributions ne sont pas identiques. Les distributions peuvent être comparées à une douche. Si elle est en position « jet étroit », presque toute l'eau est concentrée sur un seul point, c'est-à-dire le jet n'arrose pratiquement que la valeur moyenne. Si la douche est en position « pluie », l'eau est dispersée plus largement : il y a de grands écarts par rapport à la moyenne.



[1]

Pour mettre en évidence cette différence, il faut mesurer **la dispersion** des données autour de cette mesure de tendance centrale. Nous allons étudier deux mesures de dispersion.

1ère tentative *Et si nous calculions la somme des écarts à la moyenne…*

x_i	f_i	$x_i - \overline{x}$	$f_i\left(x_i - \overline{x}\right)$
165	1/9		
175	7/9		
185	1/9		
totaux			

x_i	f_i	$x_i - \overline{x}$	$f_i\left(x_i - \overline{x}\right)$
155	1/9		
165	2/9		
175	3/9		
185	2/9		
195	1/9		
totaux			

[1] Illustrations de Peter Fejes : Statistiques (les stats en bulles) / Pearson Education

La variance: La variance σ^2 d'une variable statistique est la somme (pondérée) des carrés des écarts à la moyenne:

$$\sigma^2 = f_1\left(x_1 - \overline{x}\right)^2 + f_2\left(x_2 - \overline{x}\right)^2 + \ldots + f_k\left(x_k - \overline{x}\right)^2$$

L'écart-type: L'écart-type σ est la racine carrée de la variance: $\sigma = \sqrt{\sigma^2}$

Modèle 7:

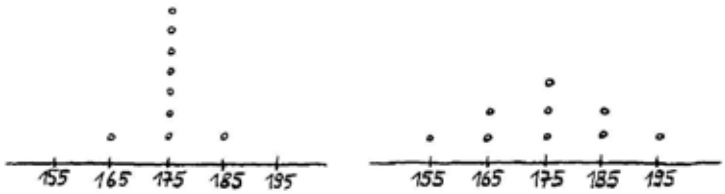
V.S quantitative discrète

Alain qui est gardien de but de l'équipe de hockey de son école, note évidemment le nombre de buts encaissés à chaque match. Il a résumé sa dernière saison dans le tableau ci-dessous (dont il s'agit de compléter quelques cases:

x_i	n_i	f_i	F_i	$f_i\,x_i$	$x_i - \overline{x}$	$f_i\left(x_i - \overline{x}\right)$	$f_i\left(x_i - \overline{x}\right)^2$
0	5	0,093	0,093	0,000	-2,611	-0,243	0,634
1	12	0,222	0,315	0,222			
2	14	0,259	0,574	0,519	-0,611	-0,158	0,097
3	8	0,148	0,722	0,444	0,389	0,058	0,023
4	7	0,130	0,852	0,519			
5	4	0,074	0,926	0,370	2,389	0,177	0,423
6	2	0,037	0,963	0,222	3,389	0,125	0,424
7	1	0,019	0,981	0,130	4,389	0,083	0,364
10	1	0,019	1	0,185	7,389	0,14	1,011
TOTAUX	54	1		2,611			

Calculer les mesures de dispersion de cette distribution.

Exercice 20.16: En reprenant la situation d'introduction:



Calculer la variance et l'écart-type

Exercice 20.17: La compagnie TEHOU a révélé les chiffres des absences de ses employés syndiqués pour deux mois d'hivers:

Nombre de jours d'absence	Nombre d'employés	
	janvier	février
0	36	17
1	35	44
2	20	38
3	11	9
4	10	7
5	2	0
6	1	0

Calculer la moyenne et l'écart-type.

Modèle 8: Le magasin de vêtements ROBA étudie depuis 90 jours ses ventes de jupes. Les données recueillies ont été regroupées en classes:

V.S quantitative continue

$[b_{i-1}\ ;\ b_i[$	x_i	n_i	f_i	F_i	$f_i\,x_i$	$(x_i-\bar{x})^2$	$f_i(x_i-\bar{x})^2$
[12 ; 16[	14	5	0,056	0,056	0,778	163,84	9,175
[16 ; 20[	18	11	0,122	0,178	2,200	77,44	9,448
[20 ; 24[	22	16	0,178	0,356	3,911	23,04	4,101
[24 ; 28[	26	21	0,233	0,589	6,067	0,64	0,149
[28 ; 32[	30	15	0,167	0,756	5,000	10,24	1,710
[32 ; 36[	34	12	0,133	0,889	4,533	51,84	6,895
[36 ; 40[	38	8	0,089	0,978	3,378	125,44	11,164
[40 ; 44[	42	2	0,022	1	0,933	231,04	5,083
TOTAUX :		90	1		26,8		47,725

Calculer les mesures de dispersion de cette distribution.

Exercice 20.18: Une étude des salaires annuels des employés d'une grande compagnie a donné les résultats suivants:

Classe	Effectifs
[20'000 ; 22'000[	80
[22'000 ; 24'000[	130
[24'000 ; 26'000[	340
[26'000 ; 28'000[	210
[28'000 ; 30'000[	120
[30'000 ; 36'000[	120

a) Calculer les mesures de tendance centrale.
b) Calculer la variance et l'écart-type.

Exercice 20.19: Lise et Michel sont deux professeurs de statistiques du gymnase. Chacun a fait passer un TE sur le sujet. Voici le tableau des notes:

Classe de Lise	1.5	2.5	3	3.5	4	4.5	5	6
	2	2.5	3.5	4	4.5	4.5	5	6
	2.5	3	3.5	4	4.5	5	6	6
Classe de Michel	2	3	3	4	4.5	5	5.5	
	2	3	3.5	4	4.5	5	6	
	2	3	3.5	4	5	5.5	6	

Déterminer la classe admettant les meilleurs résultats. S'agit-il également de celle qui est la plus homogène ?

Remarque: Vos calculatrices admettent quelques fonctions statistiques préprogrammées permettant sur une petite collection de nombres d'en calculer la moyenne et l'écart-type. N'hésitez pas à consulter votre mode d'emploi !!!

- Comparer les moyennes et les écarts-types des 3 séries de quatre notes suivantes:

  {4 ; 4 ; 4 ; 4} {2 ; 3 ; 5 ; 6} {2 ; 2 ; 6 ; 6}